Perception and Representation in a Multistrategy Learning Process

Castillo M.Dolores, Sesmero M.Paz

***Abstract*--This paper describes the features of a multistrategy learning system devoted to acquire knowledge from external data and prior knowledge. The architecture of the system is based on a perception-representation-operation cycle. In this learning context, perception process represents the capability to understand the input information and to interpret it according to the learning goals. In order to the learning fits to this model three interrelated spaces have been built that allow inference processes underlying learning strategies to work in a cooperative and competitive way.**

***Index terms*-- representation, inference, knowledge**

I. Multistrategy Learning

From a formal viewpoint, a learning process can be interpreted as a search through a knowledge space defined by the representation used. States of this space are generated by applying inference rules. The learner must be able to infer new knowledge or a better form of knowledge from input information and background knowledge and it has to record the results of inference [1]. A learning process can be defined by the input information, the background knowledge, and the goal to be achieved. The learning process determines the type of learning strategy or strategies to be used.

From a deeper understanding of the advantages and limitations of learning methods based on a sole strategy, there has been an increasing interest on multistrategy learning systems [2]. The main features of these systems are, on the one hand, the range of different complementary strategies and, on the other hand, the methods of control used to select and combine them correctly. Their versatility and power is greater. However, their complexity is also greater and so building an actual integral system presents a great challenge. Since human learning is multistrategy, research on this kind of systems has a significant relevance to understanding it, regardless of the concrete application [3].

This paper describes a multistrategy learning system based on the idea that each learning strategy is the result of certain basic inference processes which can be integrated by working within a space with a proper representation, in order to learn new knowledge or to improve existing knowledge in an iterative way. Inference processes are assembling and transforming processes of the structures in this space. The inference type characterizes knowledge changes along the truth-falsity dimension.

The input to the learning system consists of information (examples, knowledge) that the system receives from the environment. The learning goals specify criteria to be fulfilled by the learned knowledge. In a specific application of this system, this goal need to be specialized. The integration of different inference processes is intended to build knowledge units from data units by generating temporary structures which try to satisfy defined categories.

II. Perception And Representation In Machine Learning

People have the ability to analyze and understand a situation in order to make a decision in a flexible way. Getting a proper representation of data, information or knowledge involved in such a situation implies the dynamic selection of its relevant issues and their dynamic interpretation through different methods.

One of the most relevant aspects of approaching a machine learning problem is choosing a proper representation. This representation not only involves the representation formalism (rules, decision trees or neural nets) but also the attributes used in describing the examples or facts, and background and acquired knowledge, as well as the features of the learning strategies used. The objective is to get an effective representation for the overall learning process.

In this sense, the architecture underlying the multistrategy learning system is designed to simulate the steps followed when people tackle a learning problem, starting from the perception of the problem to the understanding or experience of it.

Problem representation is the result of problem perception. Perceived data (examples, facts, and knowledge) flow along a filtering and scheduling process in order to achieve a structured representation that may have multiples uses [4].

Instituto de Automatica Industrial, Ctra. Campo Real, km. 0,200. La Poveda, 28500 Arganda del Rey. Madrid. Spain.

One challenge when a learning system tries to model some aspects of human learning is understanding how to give meaning to the world: which information is relevant and the way to represent it. Problem representation must change depending on the context and on the concepts or ideas involved when the problem is perceived.

Perception is a process that activates concepts from different levels of abstraction. This ensures its flexibility since the same input data set can be perceived in different ways depending on the context and on the perceiver's state. This flexibility prevents associating a fixed representation to a perceived situation.

The proposed architecture adopts a model in which the representation is not defined beforehand but depends on the context and on the concepts involved in the learning process. This architecture works with no fixed knowledge representation but an evolving one: a structure can be represented at a given time in one way and after that time its representation can be modified as a result of the evolution of the system while dealing with the problem. That is, the learning system is able to combine perception and representation of a learning problem in such a way that its problem representation is conditioned by its problem perception and at the same time the way it perceives the problem depends on the problem representation [5].

III. TRIPARTITE ARCHITECTURE

The architecture's dynamics complies with perception-representation-operation cycles, through the interaction among three spaces, each space being responsible for one activity in a cycle: conceptual space, working space and operator space [6]. Information that every space requires is passed through defined transfer structures. The components belonging to each space can work in a parallel way and perform its interactions through these structures. The following subsections describe the main elements of every space and their function.

A. Concepts

Conceptual space represents the ideas that a human being has beforehand when tackling a learning problem. It is formed of concepts which include all the possible descriptions about perceived objects and the type of these descriptions as well as other concepts that are properties or relations among concepts. In other words, the concepts express all the notions about the objects involved in the problem and about the background knowledge needed for solving it. Conceptual space can also contain domain-dependent knowledge.

Conceptual space of this multistrategy learning system is based on the perception-representation philosophy used to solve problems by analogy [7]. According to this, concepts are characterized by their conceptual depth, degree of activation and links to another concepts.

Objects to be perceived are the initial expressions as well as the expressions generated throughout the evolution of the system working to fulfil the learning goal. These expressions are placed in the working space. Concepts must describe the features of these expressions. Conceptual depth reveals how easy it is to recognize a concept in the expressions of the current problem. The degree of activation of a concept reflects its perception within the working space.

Concepts may have inference processes associated to them. The activation of concepts promotes the work of the inference processes or operators in order to change the working space contents. The degree of activation of a concept represents the number of operations of its class that can be performed over expressions in the working space. The activation of concepts decreases when the operations associated to them have been performed.

The learning process finishes when learning goals have been achieved or when there are no activated concepts in the conceptual space.

B. Expressions

Working space is defined as a recipient of expressions (examples and knowledge) represented by logic formulas. The application of inference operators over this space allows the transformation of knowledge by creating new expressions or updating existing ones. System evolution determines that working space can contain structures with different levels of complexity.

The state of the working space at every moment reflects how the system evolves. This space can be represented by a graph made up of nodes containing the generated expressions joined by links labeled with the type of binding operator. One expression can be included in several expressions of different levels of complexity. Working space may have two classes of expressions: virtual and real expressions. Virtual expressions are generated by the action of inference operators and they are candidates to be real.

Every expression has associated to it a parameter indicating its importance within the working context regarding the depth of the concepts perceived. When an operator proposes a new expression or virtual expression to the working space, the working space decides, based on the importance of the expression, whether this expression is to be built and thus it becomes a real expression. This parameter plays an essential role by trying to prevent a disordered increase of knowledge generated by the inference processes. It represents a mechanism of dynamic weighing of the expressions, which evaluates each expression according to the system's learning goals. The building of expressions in the working space promotes the updating of concept activation within the conceptual space.

C. Operations

The integration of different learning methods requires characterizing them through common parameters in order to establish one sole representation of input information and of inference results. This architecture, instead of integrating learning strategies (induction, deduction, analogy, ...) at a macro level, performs the integration of different inference processes (specialization, generalization, abstraction, ...) which form strategies.

Inference operators are devoted to the part of operation belonging to the tripartite cycle performed by the system. The function of this space is to modify the structures of the working space according to the activated concepts of the conceptual space in order to get new knowledge or a better form of existing knowledge. This space contains different classes of inference operators. Besides, it contains operators that seek certain type of expressions and operators that compute the parameters associated to expressions in the working space. They are seeker operators and computer operators, respectively.

Every operator takes one or more expressions, according to their importance, from the working space and proposes the resulting or virtual new expression to the working space. The job of every operator is to get a better state of the working space by generating expressions that may be the solutions to the proposed learning problem. The number of instances of one operator class depends on the activation coming from the concepts of the conceptual space. Operators work in the different cycles of the learning process in the same way. It is not necessary to modify its behavior over the expressions in the working space in spite of their increasing complexity.

Every inference operator can operate regardless of the remaining ones; it is only sufficient that its application conditions be satisfied. Thus, it is possible to search new expressions in a parallel way. For instance, whereas an inductive inference process generalizes a set of examples to produce a class description, a deductive inference process can find that this description is the premise of one rule in the conceptual space.

However, it is necessary to set restrictions in the application order of the inference processes in order to ensure the very best system behavior in using resources. These criteria try to avoid the simultaneous working of unnecessary search processes, on behalf of other processes that are more efficient in the knowledge creation, or the removal by a process of the expressions previously produced by another process. Concepts, through the conceptual depth, control the order of different inference operations. The lower conceptual depth of a concept, the higher priority of its associated operators. The operators delayed become active in later learning cycles.

Most multistrategy systems have integrated learning strategies from a cooperative or competitive viewpoint. The cooperative approach is directed to the selection of the best learner whereas the competitive approach aims to get the best global result by using all the individual results.

The developed architecture fulfills both of these approaches. The system is cooperative since all the strategies contribute to a common working space and some of these strategies reuse the structures generated by other ones. It is competitive since the working space builds a real structure proposed by an operator whenever this new structure is more important than other ones previously created or those ones created at the same time by other inferential operators.

IV. Learning The Description Of A Class Of Examples

Table 1 shows a set of positive and negative examples of a class. The goal is to learn the class description from the examples provided. The examples are described by attribute-value pairs, called selectors [8].

Table 1. A sample of positive and negative examples of a class.

example	at1	at2	at3	class
1	a	x	r	+
2	a	y	r	+
3	b	y	r	+
4	b	y	s	+
5	a	y	s	-
6	a	z	r	-
7	a	z	s	-

When people approach a supervised learning problem, like the one of the table, in a manual way, they start looking at the values of the different attributes describing positive examples. Once these selectors are known, they verify whether some of these are only contained within the positive examples -the objective is to find, as soon as possible, the consistent selectors-. When consistent selectors are found, the next step is to check whether these selectors are present in all of the positive examples -thus, the expressions would then be consistent and complete-. When selectors are incomplete, it is necessary to find other consistent selectors and to perform the logic product among them (deductive specialization).

If there are no consistent selectors, the most consistent ones are chosen and the logic product is performed among them. This process is repeated until a consistent expression is found. Finally, this consistent expression has to be completed (inductive generalization).

The application of the tripartite architecture to this learning problem implies the comprehension of every term and its proper modification. Concepts must describe the features of an expression, which may be: consistent, complete, general, specific, and solution. Initial expressions are the selectors of the target class. Conceptual space has to take both of them into account by setting up the set of concepts and relations for representing them.

A few ideas underlie a supervised learning problem; they symbolize the strategy for solving it:

1. If an expression is too general, then it must be specialized
2. If an expression is too specific, then it must be generalized
3. If a consistent and complete expression is found, then this expression is a solution of the problem

Therefore, the conceptual space must contain the General, Specific and Solution concepts. The General and Specific concepts are related by unidirectional links labeled with the Complete and Consistent concepts. These relations represent the two first strategic ideas.

The Complete and Consistent concepts are easier to perceive than the Specific or General concepts and thus their conceptual depth is lower. This means that a consistent expression is perceived with greater ease or more quickly than a general expression. In the figure, selectors x and b are consistent selectors for the target class.

The Specific concept is activated whenever there exist expressions capable of being specialized, and are therefore capable of becoming consistent expressions. In the shown sample, selectors a, y, r, and s may be specialized in order to obtain consistent expressions. The same can be said about the General concept. The degree of activation of these concepts represents the number of operations of their class that can be performed over expressions in the working space. The activation of the Specific concept would be:

$$Activation = \binom{n}{2} - \binom{k}{2}$$

being n the number of expressions to be specialized and k the number of expressions whose consistency value is below a certain threshold. In this case, the activation of the Specific concept is 6.

The searching algorithm of class descriptions is an algorithm of simultaneous covering [9]: it selects the most general descriptions without removing covered examples. This algorithm chooses among alternative attribute-value pairs by comparing the subsets of data they cover. Besides, this approach allows inference operators to work in a parallel way without any synchronization. The search begins with the most general description: starting from the empty description, the following descriptions are obtained by specialization. It is a beam search that retains a defined number of candidate descriptions in every step: those with the highest consistency and completeness values. Description generation is not guided by examples of the sample. Descriptions are generated based only on the syntax of the description representation.
The Specific concept is associated with deductive and inductive specialization operators, and the General concept is associated with deductive and inductive generalization operators. The activation of the Specific concept triggers the action of specialization operators. The number of instances of deductive specialization operator is 6. Every operator instance takes two expressions (i.e.: y and r) and proposes a new or virtual expression (i.e.: $y \wedge r$).

Figure 1 shows a simplified schema of the tripartite architecture for a supervised learning problem. The different expressions of the working space are formed from simpler expressions joined by logic connectives introduced by the action of specialization and generalization operators.

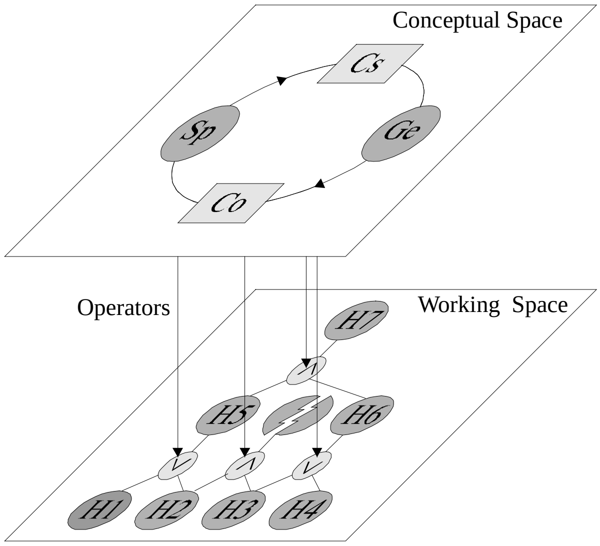

Figure 1. A partial framework of the tripartite architecture for supervised learning.

The activation of the Complete and Consistent concepts reflects the presence of expressions of these classes in the working space. These concepts do not have any associated inference processes. Their function is to allow the transfer of activation between the General and Specific concepts. The selectors b and x activate the Consistent concept and thus activation from the Specific concept is transferred to the General concept. This latter concept triggers the action of generalization operators that generate the expression $b \vee x$.

The working space evaluates the importance of created expressions and it decides whether these expressions are to be built. A consistent expression is more important than a specific expression: the consistent expression involves a concept with low conceptual depth (Consistent) and allows specialization operators to work with other less important expressions since it is impossible to keep on specializing the consistent one. The importance, for this learning goal, also considers the level of generality of a description in relation to the remaining ones. If the examples covered by a virtual description are a subset of the examples covered by a real description, the former one will not be built. If, on the contrary, the virtual description is more general than other real description, the working space will build the former and remove the latter. Figure 2 shows the evolution of the working space for the sample of Table 1 and the highlighted description is the complete and consistent description found by the system. The system has employed two cycles of perception-representation-operation for learning this description.

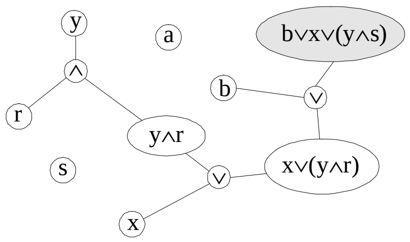

Figure 2. The description found by the learning system.

The value of the different parameters involved in this architecture are fixed in an experimental way. Their computation is out of the scope of this paper.

V. Conclusions

This paper has outlined the features of the architecture of a multistrategy learning system trying to model some aspects of human learning. The architecture of the system integrates the different inference processes in a competitive and cooperative way. Control strategy is distributed among the three parts forming the system. There is no central part governing system performance. The interaction among parts has been solved in an easy way.

The system learns by performing an infinite running cycle that consists of three stages: conceptual space, operator space and working space. When the operation of every space is small, a good simulation of their concurrence is obtained and a greater analogy with the process of perception and representation performed by people.

All of this reveals the similarity between the system and the way we think human beings solve problems of supervised learning: they use several reasoning lines and make a decision depending on the perceived conditions of the problem. If all the reasoning lines lead to the same conclusion, it is highly probable that such a conclusion be right. On the contrary, if the lines lead to contradictory conclusions and all the lines have the same power, it will be very difficult to make a conclusive decision. Although a human being could trigger his/her reasoning capabilities in a parallel way, the problem conditions determine which lines are more promising and which lines can be delayed.

VI. Acknowledgements

This work was supported by Comunidad Autónoma de Madrid with the project 07T/0028/97.

VII. Bibliography

[1] Michalski, R.S., "Machine Learning and Data Mining, Methods and Applications"., Wiley, 1998.

[2] Tecuci, G., "Plausible Justification Trees: A Framework for Deep and Dynamic Integration of Learning Strategies", Machine Learning Vol. 11, N 2/3., 1993.

[3] Michalski, R.S., "Inferential Theory of Learning as a Conceptual Basis for Multistrategy Learning", Machine Learning Vol. 11, N 2/3, 1993.

[4] Hofstadter, D., "Fluid Concepts and Creative Analogies", Basic Books, 1995

[5] Castillo, M.D., Sesmero P., "Un sistema de aprendizaje multiestratégico basado en un modelo de arquitectura fluida", Proceedings Conferencia de la Asociación Española para la Inteligencia Artificial, (pp. 72-79). Spain. 1999.

[6] Sesmero, P., del Castillo, M.D., "A Model of Three Components for the Concepts, Operations and Information Structures Involved in a Multistrategy Learning System", Proceedings of Student Sessions of Machine Learning and Applications, Grecia, 1999.

[7] Mitchell, M., "Analogy-Making as Perception. A Computer Model", MIT Press, 1993.

[8] Michalski, R.S., "A theory and methodology of inductive learning.", Michalsky R.S., Carbonell J.G., Mitchell T.M. (Eds.), Machine Learning: An Artificial Intelligence Approach., Springer-Verlag, 1983.

[9] Mitchell, T.M., "Machine Learning", McGraw-Hill, 1997.